Biological Conservation 227 (2018) 29–37

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Genetic diversity of *Ceiba pentandra* in Colombian seasonally dry tropical forest: Implications for conservation and management

Kelly T. Bocanegra-González[a,*,1], Evert Thomas[b,1], Marie-Laure Guillemin[c,d], Dulcinéia de Carvalho[e], J.P. Gutiérrez[f], C. Alcázar Caicedo[g], L.G. Moscoso Higuita[h], L.A. Becerra[f], M.A. González[i]

[a] Grupo de Investigación en Biodiversidad y Dinámica de Ecosistemas Tropicales, Universidad del Tolima, Ibagué, Colombia
[b] Bioversity International, Lima, Peru
[c] Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Valdivia, Chile
[d] CNRS, Sorbonne Universités, UPMC University Paris VI, PUC, UACH, UMI 3614, Evolutionary Biology and Ecology of Algae, Station Biologique de Roscoff, CS 90074, Place G. Teissier, 29680 Roscoff, France
[e] Laboratório de Conservação Genética de espécies Florestais, Universidade Federal de Lavras, Lavras, Brazil
[f] International Center of Tropical Agriculture (CIAT), Cali, Colombia
[g] Bioversity International, Cali, Colombia
[h] Forestpa SAS, Medellín, Colombia
[i] Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, Colombia

ARTICLE INFO

Keywords:
Genetic diversity
Microsatellites
Climate change
Seasonally dry tropical forest Paleodistribution
Suitability modelling
Conservation
Fragmentation
Seed zones

ABSTRACT

Seasonally dry tropical forests (SDTFs) are one of the most degraded vegetation types worldwide and in Colombia < 10% of the original cover remains. This calls for urgent conservation measures and restoration efforts. Understanding the genetic diversity and structure of tree species is crucial to inform not only conservation measures, but also sourcing of planting materials to ensure the long-term success of tree planting efforts, particularly in light of climate change. We assessed the genetic diversity distribution and structure of *Ceiba pentandra* from twelve representative locations of SDTF in Colombia, and how they may have been shaped by past climatic changes and human influence. We found three different genetic groups which may be the result of differentiation due to isolation of the Caribbean region, the Upper Cauca River Valley and the Patía River Valley in pre-glacial times. Range expansion of SDTF during the last glacial period, followed by more recent range contraction during the Holocene can explain the current distribution and mixture of genetic groups across contemporary STDF fragments. Most of the sampled localities showed heterozygosity scores close to Hardy–Weinberg expectations. Only two sites, among which the Patía River valley, an area with high conservation value, displayed significantly positive values of inbreeding coefficient, potentially affecting their survival and use as seed sources. While the effects of climate change might threaten *C. pentandra* populations across their current distribution ranges, opportunities remain for the *in situ* persistence of the most genetically diverse and unique ones. Based on our findings we identify priority areas for the *in situ* conservation of *C. pentandra* in Colombian SDTF and propose a pragmatic approach to guide the selection of appropriate planting material for use in restoration.

1. Introduction

Seasonally dry tropical forest (SDTF) is one of the most endangered ecosystems worldwide (Miles et al., 2006). More than 60% of its original area has already been lost, and the remaining SDTFs experience high levels of fragmentation and are little protected (Miles et al., 2006; Sánchez-Azofeifa et al., 2013). SDTFs occur in nearly all continents, but more than half (54.2%) of its surface area is found in South America where it has been an important focus for human settlement throughout history (Sánchez-Azofeifa et al., 2005; Miles et al., 2006). Colombian SDTFs occupy an intermediate position between Central and South American dry forests. They are, however, severely degraded and the

* Corresponding author.
E-mail address: ktbocanegrag@gmail.com (K.T. Bocanegra-González).
[1] K.T. Bocanegra-González and E. Thomas contributed equally to this manuscript.

https://doi.org/10.1016/j.biocon.2018.08.021
Received 1 February 2018; Received in revised form 17 July 2018; Accepted 23 August 2018
Available online 05 September 2018
0006-3207/ © 2018 Elsevier Ltd. All rights reserved.

remaining fragments are located in the Caribbean coastal area and in the Cauca, Magdalena, Chicamocha and Patía river valleys (Pennington et al., 2000; Pizano and García, 2014). These fragments contain biological elements from Caribbean and Mesoamerican SDTFs in the north and from the Andes, Brazil and southern South America in the south (Pizano and García, 2014; Banda et al., 2016). Previous research has shown that floristic differences between Caribbean and Andean SDTF in Colombia may be mirrored by the population genetic structure of tree species distributed across the different fragments (Thomas et al., 2017b).

The current genetic structure of many tropical tree species has been strongly influenced by past changes in climate. Under adverse climate conditions, a species might suffer range contraction through which spatially isolated populations undergo genetic differentiation (Thomas et al., 2015). On the other hand, more favorable climate conditions might boost range expansion and lead genetically differentiated populations to meet again in convergence zones. The two most recent climate events that are classically believed to have influenced the genetic structure of tree species from SDTF are the late Pleistocene cool and dry glacial period and climate warming during the Holocene. During the last glacial period, and especially the last glacial maximum, neotropical SDTFs are likely to have had more extensive and largely contiguous distributions (Pennington et al., 2000; Mogni et al., 2015) which may have facilitated dispersal and gene flow among species populations (Crooks and Sanjayan, 2006). Changes in climate during the warming Holocene led to contraction of SDTF eventually forming the spatially isolated fragments known today. These processes of SDTF expansion during the last glacial period and contraction in more recent times is known as the dry forest refugia hypothesis (Prado and Gibbs, 1993; Pennington et al., 2000; Mayle, 2004). Individual tree species responded to these changes in past climates in idiosyncratic ways. Evidence from paleodistribution modelling and population genetic studies in multiple SDTF tree species supports the dry forest refugia hypothesis (Caetano et al., 2008; Vitorino et al., 2016; Thomas et al., 2017b), but the opposite effect (range contraction during the last glaciation followed by range expansion) has also been reported (de Melo et al., 2016).

The post-glacial isolation of tree species populations in different SDTF fragments is likely to have initiated process of genetic differentiation. However, it seems to have been too short to be detected in the current genetic structures of SDTF tree species (but see Moreira et al., 2015). An increasing body of evidence suggests that the formation of different genetic groups in SDTF tree species may predate the late Pleistocene (Caetano et al., 2008; Collevatti et al., 2012; Vitorino et al., 2016; Thomas et al., 2017b). SDTF fragments currently dispersed throughout South America are relatively old remnants which may have persisted since the Neogene (Burnham and Carranco, 2004; Côrtes et al., 2015).

Since the arrival of humans to South America, and especially after the onset of the European colonization, all Colombian SDTF fragments have been influenced to varying extent by anthropogenic disturbance. Today, < 10% of the original SDTF cover in Colombia remains (Sánchez-Azofeifa et al., 2013; Pizano and García, 2014). Fragmentation of populations of tree species that are characterized by low densities, slow growth rates and long generation times, reduces effective population sizes, which in turn increases the risk of inbreeding and genetic drift, as documented for several species of SDTF such as *Samanea saman* (Cascante et al., 2002), *Tabebuia ochracea* (Moreira et al., 2009) and *Caryocar brasiliense* (de Melo et al., 2012). Over the longer term this might lead to loss of genetic diversity and eventually lead to extinction of local populations and concomitantly the loss of ecosystem services (Aguilar et al., 2008). Population genetic footprints of anthropogenic disturbance on populations of tree species in Colombian SDTFs are expected to be most marked in relicts located in the upper Magdalena, Chicamocha and Patía rivers due to their notable geographical isolation and associated constrained gene flow (Sarmiento, 1975; Pizano and García, 2014; Vergara, 2015). Understanding the genetic diversity and structure of tree species is crucial to inform not only conservation measures (Kageyama et al., 2003; Vieira-Blanco et al., 2007), but also sourcing of planting materials to ensure the long-term success of tree planting efforts, particularly in light of climate change (Breed et al., 2013; Thomas et al., 2014, 2017a; Prober et al., 2015).

Here we assess the genetic diversity and structure of the tree species *Ceiba pentandra* across twelve representative locations of SDTFs in Colombia in light of past climate change. In line with the findings of similar studies in tree species from neotropical STDFs, we hypothesize that the formation of genetic groups in *C. pentandra* predates the late Pleistocene and that the species experienced range expansion during the Last Glacial Maximum (LGM), leading to the admixture of different genetic groups in convergence zones. We expect LGM range expansion to have been followed by range contraction during the Holocene, resulting in an ensemble of forest fragments characterized by the presence of one or more genetic clusters. Furthermore, we investigate potential signals of human disturbance and fragmentation on the genetic profiles of sampled localities, and assess the predicted impact of climate change on their *in-situ* persistence. Based on our findings we identify priority areas for the *in situ* conservation of *C. pentandra* in Colombian SDTFs and propose a pragmatic approach to guide the selection of appropriate planting material for use in tree planting efforts.

2. Methods

2.1. Study species

Ceiba pentandra (L.) Gaertn (Malvaceae) is a pantropical habitat generalist species which is a characteristic element not only of SDTFs, but also wetter vegetation types, including tropical rainforest (Baker, 1965; Zotz and Winter, 1994). The species has become popular for its potential for reforestation of degraded areas and plywood production (Brondani et al., 2003; Celis and Jose, 2011). From an ecological point of view, *C. pentandra* fulfils an important role in protecting SDTFs soils due to the species' low nutritional requirements, high tolerance to competition and drought and its ability to regrow after disturbances (Lobo et al., 2005; Celis and Jose, 2011; Sorngmenenye-Abengmeneng et al., 2016). For instance, Paloma-Lozano (1994) showed *C. pentandra*'s higher capacity to establish in deteriorated soils of very low fertility and shallow depth compared with other native species in a SDTF in South America. However, despite these useful traits remaining *C. pentandra* populations are gradually diminishing in different parts of the world (Quesada et al., 2004; Sorngmenenye-Abengmeneng et al., 2016).

In Colombia, *C. pentandra* is present in most SDTF regions. However, due to expected negative impacts of habitat loss on the species' conservation, it has been included in the strategic plan for the conservation of plants in the country (Pizano and García, 2014). To date, there are no genetic studies of SDTF populations in Colombia to guide the development of effective genetic conservation measures, jeopardizing its sustainable management and safeguarding in the long term.

2.2. Genetic analyses

2.2.1. Field sampling

Between July 2014 and June 2015, we collected young and healthy leaves of 97 reproductive individuals of *C. pentandra* in 12 sampling sites which are representative of the current distribution of SDTF remnants in Colombia (Table 1). Except for the locality of Tatacoa (TAT), the diameter at breast height (DBH) of all trees was measured. Sampled trees were separated by at least 50 m to avoid sampling of highly genetically related individuals. Field sampling was guided by suitability maps based on occurrence data obtained from botanical collection records (see below). All biological material was collected in collaboration with the Instituto Alexander von Humboldt following the

Table 1
Genetic parameters of *Ceiba pentandra* estimated in 12 sampling sites located in Colombian SDTFs. All values are multilocus estimates based on 8 microsatellite loci.

Region	Locality	ID	*N*	A_O	A_{rich}	% PA	Ho	He	F_{is}	F_{st}	F_{st} (SE)
Caribbean Coast	Colosó	COL	8	6	3.98	0.00	0.72	0.77	0.13*	0.039	0.018
	El Guamo	GUA	14	7.38	4.01	0.00	0.92	0.80	−0.09*	0.044	0.019
	Zambrano	ZAM	11	7.13	4.00	0.00	0.88	0.79	−0.06	0.044	0.017
Cauca River Valley	Palmira	CIAT	3	4.00	4.00	0	0.96	0.70	–	–	–
	Santa Fe de Antioquia – Cotové	SFE-COT	10	8	4.51	1.83	0.90	0.84	−0.02	0.041	0.018
	Mata Guadua	MAT	4	4	3.89	1.83	0.91	0.72	–	–	–
	La Paila	PAI	12	6	3.71	0	0.78	0.75	0	0.062	0.018
	La Pintada	PIN	11	7	4.27	3.66	0.83	0.82	0.03	0.036	0.015
Magdalena River Valley	Tatacoa	TAT	3	4	4.50	1.83	0.96	0.74	–	–	–
	Tolima	TOL	2	2	–	–	–	–	–	–	–
Patía River Valley	Patía	PAT	9	4	3.00	4.58	0.48	0.61	0.26*	0.149	0.006
Chicamocha Canyon	Chicamocha	CHI	10	8	4.30	1.83	0.83	0.82	0.03	0.034	0.015

ID abbreviation of sampling sites; *N* sample size; A_O average number of alleles per locus; A_{rich} allelic richness with correction of sample size; %PA percent of private alleles; Ho observed heterozygosity; He expected heterozygosity, F_{is} multilocus estimate of the inbreeding coefficient; F_{st} multilocus estimate of the genetic differentiation with the rest of the sites sampled.

* Significantly different from 0 at $p > 0.05$.

Colombian Decree 302 115 of 2003.

2.2.2. DNA extraction and PCR amplification

We extracted total genomic DNA from dry leaves using the Doyle and Doyle (1987) protocol with modifications described by Novaes et al. (2009). Polymerase chain reactions (PCR) were carried out for 8 microsatellite markers developed by Brondani et al. (2003) for *C. pentandra*. Procedure for fluorescent dye labelling of PCR fragments was based on Schuelke (2000). The reaction was performed in a total volume of 12 μL containing 0.5 U of Platinum® Taq (Invitrogen®, USA), 1× PCR Buffer (200 mM Tris-HCl (pH 8.4)), 500 mM KCl (Invitrogen®, USA), 0.2 mM of each dNTP (Promega Corp., USA), 8.0 mg/mL BSA, 3.0 mM MgCl2 (Invitrogen®, USA), 0.05 pmol/μL Labelled M13 primer, 0.01 pmol/μL M13 tagged-Forward primer, 0.05 pmol/μL reverse primer and 40 ng of DNA. PCR was conducted following Brondani et al. (2003). PCR products were run on an ABI PRISM 3730 DNA Analyzers sequencer and sized with GeneScan-500LIZ (Applied Biosystems) size standard. Allele sizes were determined using GeneMapper version 4.0 (Applied Biosystems) software.

2.2.3. Data analyses

For each *C. pentandra* sampling site we calculated the average number of alleles per locus (Ao), the percentage of private alleles, the observed (Ho) and expected (He) heterozygosity under Hardy–Weinberg equilibrium (Nei, 1978) and the allelic richness (A_{rich}) using GenAlEx 6.5 and FSTAT v.2.9 softwares (Peakall and Smouse, 2012; Goudet, 1999). The presence of null alleles was tested using the Micro checker software, the frequency of null allele for each locus was estimated using the Brookfield equation: $r = (He - Ho) / (1 - He)$. (Van Oosterhout et al., 2004).

Deviations from Hardy–Weinberg proportions were quantified by the unbiased estimator of Wright's inbreeding coefficient F_{IS} calculated according to Weir and Cockerham (1984). To analyze differentiation among-localities we considered each sampling site as a distinct population and calculated levels of differentiation using pairwise estimates of F_{ST} following Weir and Cockerham (1984). Both analyses were carried out in FSTAT v.2.9 (Goudet, 1999). To investigate possible isolation by distance (IBD) patterns we carried out Mantel tests in GENETIX (Belkhir et al., 1996–2004) using the correlation between $F_{ST}/(1 - F_{ST})$ estimates and the natural logarithm of the geographical distances between population pairs as proposed by Rousset (1997).

We carried out Bayesian cluster analysis with STRUCTURE 2.2 (Pritchard et al., 2000). Analyses were performed using an admixture model with 500,000 iterations and a burn-in of 100,000 iterations. Ten replicates for each K were analyzed with K = 1–13. To visualize the estimated likelihood of each run, we calculated the coefficient of similarity between runs and Evanno's ΔK (Evanno et al., 2005). We used CLUMPP (Jakobsson and Rosenberg, 2007) and DISTRUCT (Rosenberg, 2004) to combine assignment of clusters across replications and visualize the output, respectively. We mapped the proportions of assignment of the localities to the different clusters both for all individuals per population and only those individuals with high cluster membership. Cluster membership thresholds are often used is population genetics for assigning individuals to genetic groups and to filter out admixed individuals. Thresholds as low as 0.5 up to 0.8 have been used in literature (Motamayor et al., 2008; Larranaga et al., 2017; Alves-Pereira et al., 2018) and here we have opted to use an intermediate value of 0.65. As individuals with high cluster membership are less likely to derive from pollen-mediated gene flow, they can prove more useful for unravelling the potential origin of genetic groups, as well as the direction and genetic implications of range dynamics.

2.3. Suitability modelling

We characterized the spatial distribution of favorable habitat for *C. pentandra* in Colombia under different climatic conditions by means of suitability mapping based on ensembles of modelling algorithms, implemented in R package *BiodiversityR* (Kindt and Coe, 2005). Habitat suitability during the LGM (~21,000 BP) and Mid Holocene (~6,000 BP) was modeled to determine whether there is evidence that past climate conditions had an impact on the current distribution of genetic diversity. Habitat suitability under present and future climate conditions was modeled to assess the expected impact of climate change on the effectiveness of *in situ* conservation of *C. pentandra* populations.

Presence data collected during field sampling were complemented with Colombian records extracted from numerous sources (www.gbif.org; the national herbaria MEDEM, HUA y MEDEL; www.dryflor.info; www.orinoquiabiodiversa.org; www.sibcolombia.net). Only records from SDTF as defined by the combination of Etter et al. (2008) and García et al. (2014) were included in our dataset. As a result, 228 presence points were used for suitability modelling. Background points (an overall maximum of 10,000 and maximum one per grid cell) were randomly selected from the area enclosed by a convex hull polygon constructed around all presence points and extended with a buffer corresponding to 10% of the polygon's largest axis.

We applied two different strategies for suitability modelling under past and future climate conditions. Model calibrations for projections to LGM and mid-Holocene climate conditions were carried out at 30 arc seconds resolution using only WorldClim climate layers (Hijmans et al., 2005) as explanatory variables. Model calibrations intended for

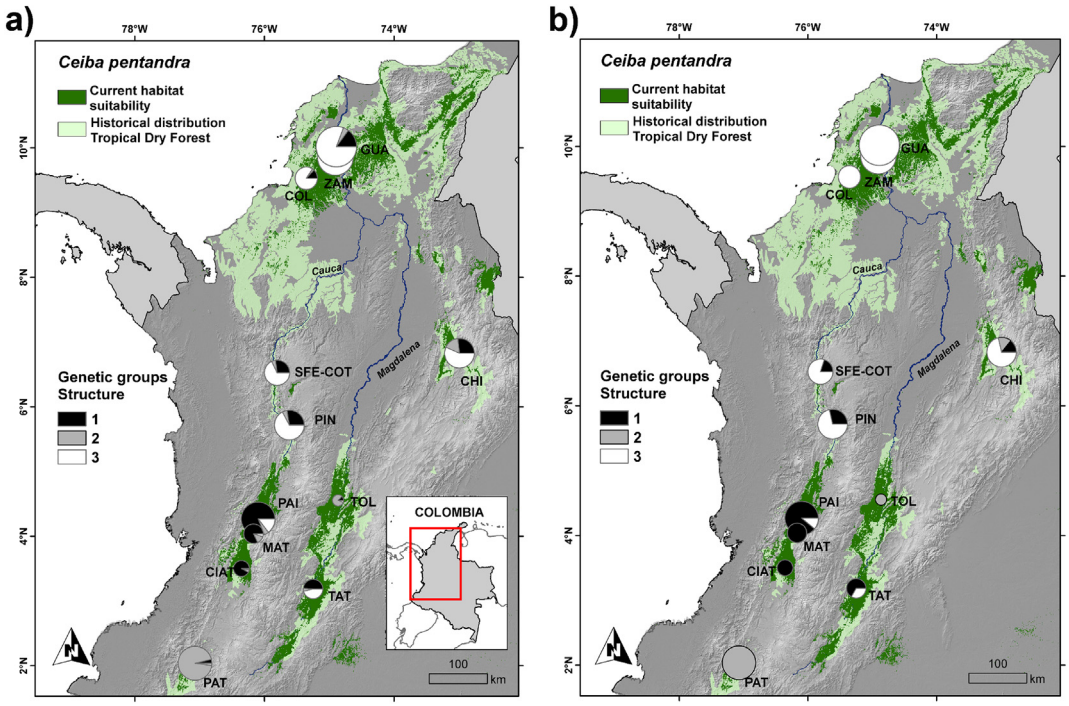

Fig. 1. Distribution of genetic groups across sampled localities of *Ceiba pentandra*. a. Cluster memberships averaged over all sampled individuals per locality; b. Numbers of individuals assigned to different clusters based on membership values of 0.65 or higher. Circles sizes are proportionate to number of individuals.

projections to future climate scenarios (period 2040–2069; referred to as 2050s) were carried out at 30 arc seconds resolution, using aside from climate layers also altitude, slope, aspect, terrain roughness, direction of water flow and seven major edaphic variables, obtained from ISRIC-World Soil Information (Hengl et al., 2014): organic carbon (ORCDRC), pH in H2O (PHIHOX), sand % (SNDPPT), silt % (SLTPPT), clay % (CLYPPT), Cation Exchange Capacity (CEC), Bulk density (BLD), Coarse fragments > 2 mm (CRFVOL). For the edaphic variables we calculated a weighted mean across 0–5, 5–15, 15–30, 30–60, and 60–100 cm soil depth values in order to derive a single data value for 0–100 cm. Collinear explanatory variables were removed based on iterative calculations of variance inflation factors (VIF), retaining only variables with VIFs smaller than 5. The resulting sets of explanatory variables, as well as presence, background and absence points used for model calibrations are given in Suppl. Table S1.

Modelling algorithms considered in the ensembles were maximum entropy (MAXENT), boosted regression trees (BRT, including a stepwise implementation), random forests (RF), generalized linear models (GLM; including stepwise selection of explanatory variables), generalized additive models (GAM; including stepwise selection of explanatory variables), multivariate adaptive regression splines (MARS), regression trees (RT), artificial neural networks (ANN), flexible discriminant analysis (FDA), support vector machines (SVM), and the BIOCLIM algorithm. As spatial autocorrelation among species presence points is known to bias model evaluations based on cross-validation, we evaluated the ability of all individual modelling algorithms to cope with spatial autocorrelation by calculating calibrated Area Under Curve (cAUC) values and comparing these with a geographical null model (Hijmans, 2012). We compared the cAUCs of each of the individual distribution models with the cAUCs of the geographical null model resulting from twenty iterations, by means of Mann-Whitney tests. Only models that gave cAUC values that were significantly higher than the null model were retained for the construction of different model ensembles. In a next step, we calculated both the calibrated AUC values for all possible ensemble combinations of the retained models. Each ensemble combination was constructed as the weighted average of its individual composing models, using the cAUC values as weights. The ensemble that yielded the highest cAUC value was considered to generate the most appropriate scenario for projecting to past and future climate conditions, respectively.

To assess habitat suitability under mid-Holocene and LGM climate conditions we carried out projections to two climate models each (BCC-CSM1-1 and CCSM4, and MIROC-ESM and CCSM4, respectively). For characterizing future climate conditions, we used 30 downscaled climate models for the period 2040–2069 based on the Representative concentration pathway (RCP) 4.5 scenario of greenhouse gas emissions, prepared for the Fifth Assessment IPCC report (CMIP5) (Ramirez Villegas and Jarvis, 2010). We limited model projections to areas where suitability scores were higher than the maximum training sensitivity plus specificity threshold obtained from model calibration under current climate conditions. To obtain summarizing maps for the two LGM and mid-Holocene climate models we averaged the threshold-limited suitability maps constructed for both individual climate scenarios. Two scenarios were considered for future suitability maps (period 2040–2069, referred to as the 2050s). Optimistic and pessimistic scenario maps were limited to areas which were identified as suitable by at least one and half of all 30 possible threshold-limited climate projections, respectively. Model statistics are given in Suppl. Table S1.

3. Results

The microsatellites used yielded moderately to highly variable allele numbers per locus, ranging from 9 for locus L7 to 19 for locus L3 (locus names are as reported in Brondani et al., 2003). The Patía (PAT) site was least diverse (A_{rich} = 3.00 and expected heterozygosity He = 0.61) of all sites sampled (Table 1). The highest genetic diversity was recorded for sites located in the Middle Cauca River Valley (SFE-COT, PIN), the Chicamocha Canyon (CHI) and the Upper Magdalena River Valley (TAT) (A_{rich} > 4.25 and He > 0.74; Table 1). After Patía (% PA = 4.58), these were also the localities holding the highest proportions of private alleles (% PA from 1.83 to 3.66) (Table 1). Inbreeding coefficient varied from −0.09 to 0.26, but only the values obtained in El Guamo (GUA, Fis = −0.09), Colosó (COL; Fis = 0.13) and Patía (PAT, Fis = 0.26) differed significantly from Hardy-Weinberg

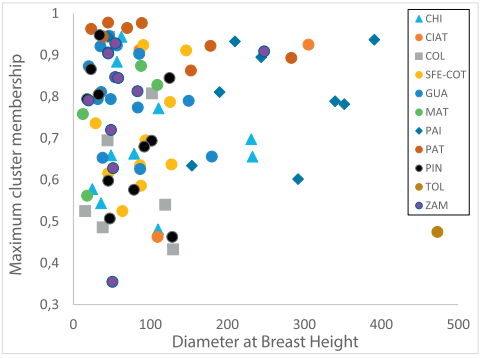

Fig. 2. Diameter at breast height of *Ceiba pentandra* trees plotted against their maximum cluster membership.

expectations ($p < 0.05$; Table 1). Suppl. Fig. S1 shows the spatial distribution of the genetic diversity parameters against a background of the species current habitat suitability and the historical distribution of SDTF in Colombia.

Results of ΔK computation showed support for K = 3 for the different scenarios modelling tested in STUCTURE (Suppl. Fig. S2a). All sampling sites, except Patía (PAT), showed signs of admixture (Fig. 1a and Suppl. Fig. S2b). However, when only individuals were retained with cluster membership of 0.65 or higher, more coherent geographical patterns emerged about the potential origin of clusters (Fig. 1b). One cluster was predominant in the Caribbean region (white cluster in Fig. 1b; COL, GUA, ZAM), another in the Upper Cauca River Valleys (black cluster in Fig. 1b; PAI; MAT, CIAT) and the third one was mostly restricted to the Patía River Valley (grey cluster in Fig. 1b; PAT). Most trees of mixed origin, with a maximum cluster membership of 0.65 or less, had diameters at breast height (DBH) below 150 cm (Fig. 2) with the notable exception of a tree with DBH of 473 cm from Tolima (TOL, Magdalena River Valley) which received equal membership scores for the grey and black genetic clusters (Fig. S2b).

A significant IBD pattern was observed when all localities were considered (Mantel test; $p < 0.01$; Suppl. Fig. S3). However, this pattern was due to the genetic isolation of the Patía (PAT) site and the pattern disappeared when PAT was excluded.

The modeled distributions of suitable habitat during past climates suggest that *C. pentandra* populations from SDTF in Colombia may have had a much wider range during the LGM than at present (Fig. 3a) and that current agglomerated distribution patterns may have started to form in the mid-Holocene (Fig. 3b). The LGM map suggests that suitable habitat conditions were found in a large continuous area connecting the Caribbean region with the Chicamocha Canyon and Magdalena River valleys, permitting south- and eastward range expansion of the genetic cluster with putative origin in the Caribbean region (white cluster; Fig. 3a). The map furthermore suggests that the Patía and upper Cauca River valleys may have been connected by suitable habitat conditions with the Magdalena River valley, permitting the expansion of the black and grey genetic clusters into the latter valley Fig. 3a (full grey and black arrows). Interestingly, according to our models, the middle Cauca River valleys were unsuitable for *C. pentandra* during the mid-Holocene, implying that the current populations found there are the result of more recent range expansion from southern and northern origins, respectively (striped arrows in Fig. 3a).

Future climate projections suggest that the impact of climate change on *C. pentandra* habitat suitability might be considerable (Fig. 4). While nearly all currently suitable sites are likely to remain so and might even expand in the near future according to at least one of the 31 climate models considered (optimistic scenario; Fig. 4a), vast areas in the Caribbean region and Magdalena River Valleys are expected to become unsuitable according to at least half of all future climate model projections (pessimistic scenario; Fig. 4b). Even under the latter scenario, extensive suitable areas remain in all areas we sampled, suggesting that local populations and the genetic diversity they contain might be able to persist in the near future under climate change. Except from the *los Colorados fauna and flora sanctuary* (small black dot above the "M" of "ZAM" in Fig. 4) in the Caribbean region, one of the official protected areas of SDTF (Suppl. Table S2).

4. Discussion

Here we assessed the genetic diversity distribution of *C. pentandra* localities in Colombian SDTF and how it may have been shaped by past climatic changes and more recent human influences. Our LGM and mid-Holocene habitat suitability models for *C. pentandra* are consistent with the dry forest refugia hypothesis (Prado and Gibbs, 1993; Pennington et al., 2000; Mayle, 2004). Joint interpretation of this finding with the current distribution of the three genetic groups we found, suggests that the genetic differentiation of the latter is likely to predate the LGM, in line with similar findings for other species (Caetano et al., 2008; Collevatti et al., 2012; Vitorino et al., 2016; Thomas et al., 2017b). During the LGM, genetic clusters of *C. pentandra* with putative origin in the upper Cauca and the Patía River valleys, respectively, are likely to have been partly connected by suitable habitat conditions across the central Andes chain with the Magdalena River valleys, permitting their migration into the latter valleys (full grey and black arrows Fig. 3a). *Ceiba pentandra* currently occurs from sea level up to 2000 m.a.s.l. (Fernández-Alonso, 2015) and during the LGM this upper boundary may have been higher, as suggested by our LGM model which identified several areas in the central Andes chain above 3000 m.a.s.l. as suitable. Similarly, suitable habitat conditions in a large continuous area connecting the Caribbean region with the Chicamocha Canyon and Magdalena River valleys, might have facilitated southward migration of the genetic group which is likely to have originated in the Caribbean region (full white arrows Fig. 3a).

Our suitability models furthermore suggest that habitat conditions are likely to have remained stable in many of our sampling sites since the LGM, or possibly longer. Notable exceptions are the localities from the Caribbean region (COL, ZAM, GUA) and the middle Cauca River valley (SFE-COT, PIN). In the Caribbean region, suitable habitats may have been more fragmented during the mid-Holocene, but stable areas remained in the vicinities of our sampling locations. By contrast, most of the middle Cauca River valley seemed to have been unsuitable during the mid-Holocene. Areas where habitat conditions have remained stable over prolonged periods of time may often display higher levels of genetic diversity, while unstable areas are expected to have been colonized only more recently and display overall lower or shallower levels of diversity (Hewitt, 2004; Carnaval et al., 2009; Thomas et al., 2015, 2017c). Contrary to these expectations, nearly all of the sampled localities, including those from the middle Cauca River valley, yielded similarly high genetic diversity scores, regardless of habitat stability. Furthermore, only the population from the Patía River Valley region, where habitat suitability is likely to have remained stable, was markedly less diverse and was the genetically most isolated population of all, as suggested by the highest genetic distances and numbers of private alleles documented for this population.

We postulate that the diversity scores we found in the Patía River Valley region and middle Cauca River valleys to be lower and higher than expected, respectively, are related to processes of genetic isolation and anthropogenic degradation in the Patía region and the fact that the middle Cauca River valley is a convergence zone of genetic clusters of different origin. While the positive correlation between habitat stability and genetic diversity may be true for large population sizes in pristine vegetation, it is likely to be obscured in areas that have suffered the impacts of longstanding anthropogenic disturbance, and/or have been isolated for prolonged periods of time. Fragmentation is known to negatively affect the reproduction, gene flow and genetic diversity of tree populations (Aguilar et al., 2006, 2008; Moreira et al., 2015). The Patía

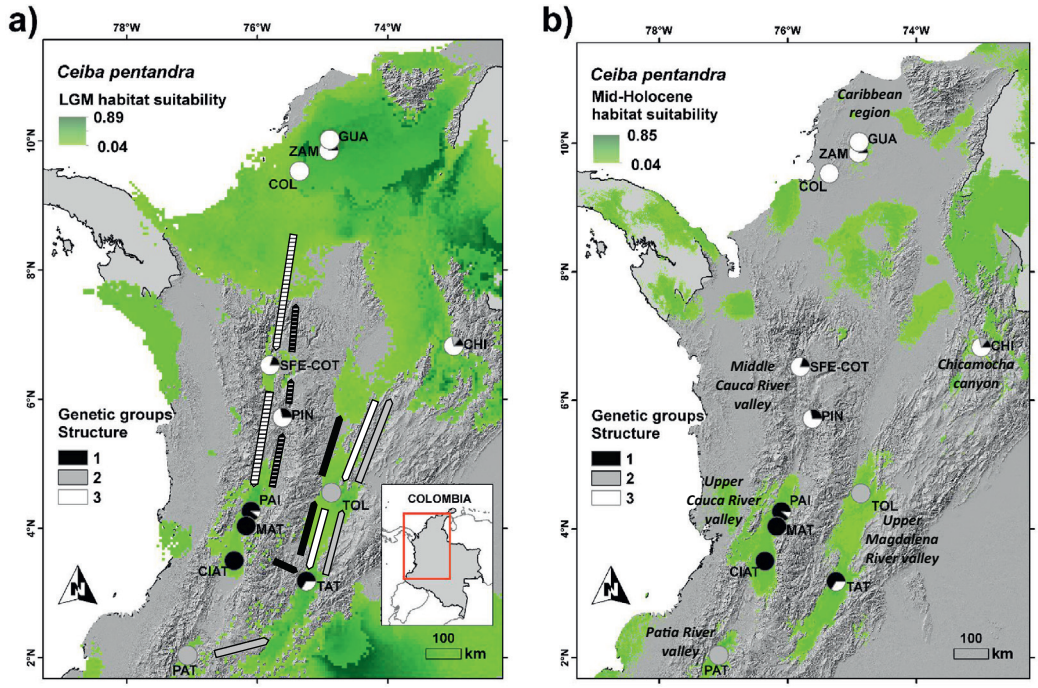

Fig. 3. Distribution of suitable habitat conditions for *Ceiba pentandra* during the LGM (a) and mid-Holocene (b). Genetic groups correspond with those in Fig. 1b, arrows indicate potential directions of range expansion during the last glacial period (full arrows) and the period following the mid-Holocene (striped arrows). Arrows colors correspond to genetic groups.

(PAT) locality yielded both the highest values of the inbreeding coefficient and the lowest values of genetic diversity of all studied sites. Lower genetic diversity of this locality might be the result of prolonged isolation, a very low influx of genes from external sources, and possibly, the effects of genetic drift due to a smaller effective population size. More recent anthropogenic fragmentation of Patía forests due to conversion to pasture and crop land and mining activities (Vergara, 2015) is likely to have further aggravated this situation, explaining the significant positive inbreeding coefficient. In degraded landscapes, visitation of *Ceiba* flowers by its main bat dispersers is lower (Quesada et al., 2003) and trees tend to self-pollinate more than in undisturbed areas (Lobo et al., 2005).

Interestingly, the highest diversity scores were consistently observed in localities with genotypes of mixed origin (SFE-COT, PIN, TAT, CHI). This is in concordance with previous findings that recent colonization areas may hold high genetic diversity in convergence zones where genetically differentiated populations have come in contact (Petit et al., 2003; Marchelli et al., 2017). Probably, the best example of

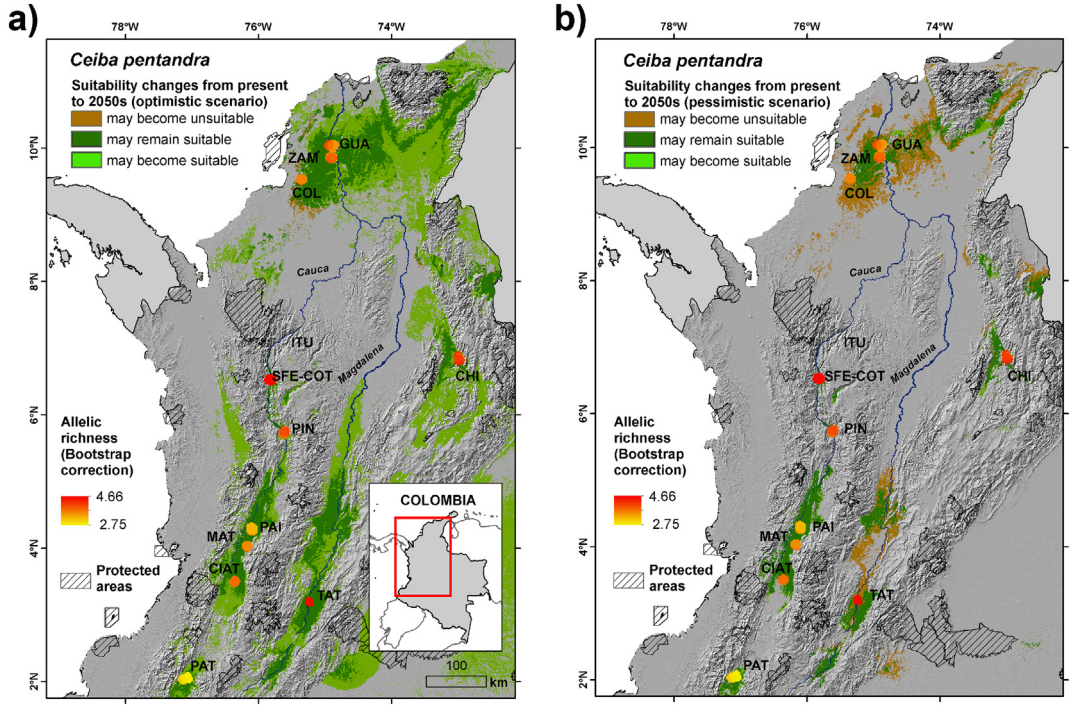

Fig. 4. Future (2050s) habitat suitability of *Ceiba pentandra* showing all areas identified as suitable by at least 1 (a) and at least 15 (b) of 31 future climate model projections, respectively, in combination with the distribution of allelic richness. Black stripped areas represent current Colombian protected areas; please note that the Colorados sanctuary in the Caribbean region correspond to the small black dot above the M of "ZAM".

this are the middle Cauca River valley localities which are likely to have formed most recently (after the mid-Holocene) through range expansion of populations from southern and northern origin, respectively (Fig. 3a). Similarly, elevated diversity observed in the SDTF trees *Enterolobium cyclocarpum* and *Anacardium excelsum* at the northern end of the middle Cauca River Valley is thought to be the result of mixing of genotypes from southern and northern origins (Thomas et al., 2017b; Bocanegra-González and Guillemin, 2018). By contrast, the mixed origin of upper Magdalena and Chicamocha River localities might date back to the LGM, or longer.

The absence of signs of inbreeding in nearly all localities except Patía (PAT), despite centuries of human fragmentation, might be due to the effects of historical and ongoing gene flow across populations, evident from the admixed nature of all populations (Fig. S2b) and low degree of genetic differentiation among them (Fig. S3). *Ceiba pentandra* is capable of long distance gene flow, mainly though anemochorous seed dispersal (Dick et al., 2007). Long distance dispersal of pollen can also play a role since pollinating bats have foraging distances up to several kilometers (Lobo et al., 2005). Furthermore, our data suggest that cross-population gene flow might have increased in recent times. Indeed, nearly all the genotyped individuals with maximum cluster membership below 0.65 had diameters below 150 cm (Fig. 2), implying that they probably established around the beginning of the past century (https://www.monumentaltrees.com/en/trees/ceibapentandra/records/). As this coincides with the intensification of human land transformation (Vina and Cavelier, 1999), it could point to human-mediated dispersal of *C. pentandra* seeds or seedlings from different origins.

The evidence of inbreeding we found in the Colosó (COL) population was unexpected, as this was believed to be one of the least degraded sites we visited, based on visual inspection of vegetation (*e.g.* closed canopy, presence of indicator species such as large-sized *Swietenia macrophylla* trees which are typically the first ones to be logged). Inbreeding is one of the early signs of habitat fragmentation, while genetic diversity is lost more slowly over subsequent generations (Lowe et al., 2015). Interestingly, genetic diversity in Colosó was as high as in the other Caribbean localities and all trees sampled had diameters below 130 cm. This might point to the effects of a reduction in effective population size, possibly due to some form of anthropogenic disturbance that currently goes unnoticed in rapid visual assessments of the conservation state of local vegetation. A positive value of inbreeding coefficient was also found in Colosó for the *E. cyclocarpum* (Thomas et al., 2017b), further supporting this hypothesis. While further research is needed to elucidate the nature of potential anthropogenic impacts on tree populations in Colosó, our findings suggest that assessing the conservation state of a forest through visual inspection may be misleading and needs to be validated with genetic tools.

5. Implications for conservation and management

The recent commitments of Latin American countries to restore tens of millions of hectares of degraded lands on their territories, in the context of Initiative 20 × 20 (http://www.wri.org/our-work/project/initiative-20x20), provide opportunities to reverse the trend of forest degradation and fragmentation and improve the prospects for conservation of the remaining populations of *C. pentandra*. Due to the species' hardiness to survive under a variety of harsh habitat conditions, its use in forest and landscape restoration projects is bound to increase. While a growing body of knowledge on the sexual and asexual propagation and management of the species (Khurana and Singh, 2001; Celis and Jose, 2011) exist, limited attention has thus far been given to the genetic quality of planting material. Ensuring that reproductive material of *C. pentandra* used in tree planting efforts will survive, thrive, and establish viable and resilient populations in the long term requires (1) adequate planning to identify the seed sources that are best matched to the conditions of the planting site and (2) application of good seed collection protocols (Thomas et al., 2014). The selection of germplasm should ideally be guided by the strength of the interaction between genotype performance and current and expected future environmental conditions (genotype-by-environment interactions), which are studied using multilocation progeny or provenance trials and climate modelling, respectively (Sgro et al., 2011; Breed et al., 2013). However, to the best of our knowledge, no provenance trials existed for *C. pentandra* in Colombian SDTFs, and at the time of this publication the first progeny trial was only in the initial phase of establishment by some of the authors. In the absence of multi-year observations from such trials, a combination of the neutral genetic characterization data and climate modelling, similar to the ones we have reported here can be used as a best guess approach (Thomas et al., 2014).

Although most populations we sampled contained genotypes of mixed origin, it seems appropriate to respect the putative origin of the three genetic clusters we identified in seed sourcing activities. In other words, tree planting activities in the Caribbean region, the upper Cauca River valleys and the Patía River valleys should preferentially use reproductive material collected from local to regional sources. This recommendation may later be adjusted based on observations from field trials or complementary genetic characterization studies. In areas where genotypes of more than one cluster occur it might be best to source material from the same region (*e.g.* middle Cauca and Magdalena River valleys, Chicamocha canyon). For ensuring genetic diversity in planting stock, source populations should be large (ideally at least 500 reproductively mature individuals), and seeds should be obtained from a high number (> 15) of mother trees per population, among a series of other considerations recently summarized by Basey et al. (2015). Healthy source populations should be identified in each of the regions where the three clusters are likely to have originated and be the focus of *in situ* conservation measures. Such populations should best have high genetic diversity, show absence of inbreeding and be located in areas where habitat conditions are expected to remain suitable in the future. Reproductive material collected in Colosó, which scored high for the inbreeding coefficient, should be mixed with materials from nearby localities within the Caribbean region. As the contribution of a population to total species diversity depends more on its divergence from other populations than on its intrinsic within-population diversity (Petit et al., 2003), the genetic uniqueness of the Patía (PAT) locality should largely outweigh its low diversity for long-term conservation purposes. The establishment of the Patía Fauna and Flora Sanctuary conservation unit proposed by the Colombian national system of natural parks (PNNC, 2018) should therefore be implemented without further delays. However, seed collection efforts for tree planting in this region should aim for the collection of seeds from a large number of mother trees (> 30).

Acknowledgments

The authors wish to thank the Colombian companies Ecopetrol and Empresas Publicas de Medellin, the Government of the Colombian department of Antioquia, the CGIAR Fund Donors (http://www.cgiar.org/who-we-are/cgiar-fund/fund-donors-2) and the CGIAR research program on Forest Trees and Agroforestry for financial support. This research was further supported by the Organization of American States (OAS), Conselho Nacional de Desenvolvimento Científico e Tecnológico of Brazil (CNPq).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biocon.2018.08.021.

References

Aguilar, R., Ashworth, L., Galetto, L., Aizen, M.A., 2006. Plant reproductive susceptibility

to habitat fragmentation: review and synthesis through a meta-analysis. Ecol. Lett. 9, 968–980.
Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y., Lobo, J., 2008. Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. Mol. Ecol. 17, 5177–5188.
Alves-Pereira, A., Clement, C.R., Picanço-Rodrigues, D., Veasey, E.A., Dequigiovanni, G., Ramos, S.L.F., Pinheiro, J.B., Zucchi, M.I., 2018. Patterns of nuclear and chloroplast genetic diversity and structure of manioc along major Brazilian Amazonian rivers. Ann. Bot. 1–15.
Baker, H.G., 1965. The evolution of the cultivated kapok tree: a probable West African product. In: Brokesnsha, D. (Ed.), Ecology and Economic Development in Africa. Institute of International Studies/University of California, Berkeley, pp. 185–216 chap. 9. (Research series, n. 9).
Banda, K., Delgado-Salinas, A., Dexter, K.G., Linares-Palomino, R., Oliveira-Filho, A., Prado, D., Pullan, M., Quintana, C., Riina, R., Rodríguez, G.M., Weintritt, J., Acevedo-Rodríguez, P., Adarve, J., Álvarez, E., Aranguren, A., Arteaga, J.C., Aymard, G., Castaño, A., Ceballos-Mago, N., Cogollo, C., Cuadros, H., Delgado, F., Devia, W., Dueñas, H., Fajardo, L., Fernández, A., Fernández, M.A., Franklin, J., Freid, E.H., Galetti, L.A., Gonto, R., González-M, R., Graveson, R., Helmer, E.H., Idárraga, A., López, R., Marcano-Vega, H., Martínez, O.G., Maturo, H.M., McDonald, M., McLaren, K., Melo, O., Mijares, F., Mogni, V., Molina, D., Moreno, N.P., Nassar, J.M., Neves, D.M., Oakley, L.J., Oatham, M., Olvera-Luna, A.R., Pezzini, F.F., Reyes Dominguez, O.J., Ríos, M.E., Rivera, O., Rodríguez, N., Rojas, A., Särkinen, T., Sánchez, R., Smith, M., Vargas, C., Villanueva, B., Pennington, R.T., 2016. Plant diversity patterns in neotropical dry forests and their conservation implications. Science 353, 1383–1387.
Basey, A.C., Fant, J.B., Kramer, A.T., 2015. Producing native plant materials for restoration: 10 rules to collect and maintain genetic diversity. Nativ. Plants J. 16, 37–53.
Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F., 1996–2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Genome, Populations, Interactions, CNRS UMR 5000. Université de Montpellier II, Montpellier, France URL. http://www.genetix.univmontp2.fr/genetix/intro.htm.
Bocanegra-González, K.T., Guillemin, M.-L., 2018. Guidelines for the restoration of the tropical timber tree *Anacardium excelsum*: first input from genetic data. Tree Genet. Genomes 14, 59.
Breed, M.F., Stead, M.G., Ottewell, K.M., Gardner, M.G., Lowe, A.J., 2013. Which provenance and where? Seed sourcing strategies for revegetation in a changing environment. Conserv. Genet. 14, 1–10.
Brondani, R.P.V., Gaiotto, F.A., Missiaggia, A.A., Kirst, M., Gribel, R., Grattapaglia, D., 2003. Microsatellite markers for *Ceiba pentandra* (Bombacaceae), an endangered tree species of the Amazon forest. Mol. Ecol. Notes 3, 177–179.
Burnham, R.J., Carranco, N.L., 2004. Miocene winged fruits of *Loxopterygium* (Anacardiaceae) from the Ecuadorian Andes. Am. J. Bot. 91, 1767–1773.
Caetano, S., Prado, D., Pennington, R.T., Beck, S., Oliveira-Filho, A., Spichiger, R., Naciri, Y., 2008. The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). Mol. Ecol. 17, 3147–3159.
Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T., Moritz, C., 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. Science 323, 785–789.
Cascante, A., Quesada, M., Lobo, J.J., Fuchs, E.A., 2002. Effects of dry tropical forest fragmentation on the reproductive success and genetic structure of the tree *Samanea saman*. Conserv. Biol. 16, 137–147.
Celis, G., Jose, S., 2011. Restoring abandoned pasture land with native tree species in Costa Rica: effects of exotic grass competition and light. For. Ecol. Manag. 261, 1598–1604.
Collevatti, R.G., Terribile, L.C., Lima-Ribeiro, M.S., Nabout, J.C., De Oliveira, G., Rangel, T.F., Rabelo, S.R., Diniz-Filho, J.A.F., 2012. A coupled phylogeographical and species distribution modelling approach recovers the demographical history of a Neotropical seasonally dry forest tree species. Mol. Ecol. 21, 5845–5863.
Côrtes, A.L.A., Rapini, A., Daniel, T.F., 2015. The *Tetramerium lineage* (Acanthaceae: Justicieae) does not support the Pleistocene Arc hypothesis for South American seasonally dry forests. Am. J. Bot. 102, 992–1007.
Crooks, K.R., Sanjayan, M., 2006. Connectivity conservation: maintaining connections for nature. Conserv. Biol. 14, 1.
Dick, C.W., Bermingham, E., Lemes, M.R., Gribel, R., 2007. Extreme long-distance dispersal of the lowland tropical rainforest tree *Ceiba pentandra* L. (Malvaceae) in Africa and the Neotropics. Mol. Ecol. 16, 3039–3049.
Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.
Etter, A., McAlpine, C., Possingham, H., 2008. Historical patterns and drivers of landscape change in Colombia since 1500: a regionalized spatial approach. Ann. Assoc. Am. Geogr. 98, 2–23.
Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620.
Fernández-Alonso, J.L., 2015. *Ceiba pentandra* (L.) Gaertn. In: Bernal, R., Gradstein, S.R., Celis, M. (Eds.), Catálogo de plantas y líquenes de Colombia. Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogotá. http://catalogoplantasdecolombia.unal.edu.co (retrived 23 January 2018).
García, H., Corzo, G., Isaacs, P., Etter, A., 2014. Distribución y estado actual de los remanentes del bioia: insumos para su gestiónma de bosque seco tropical en Colomb. In: Pizano, C., García, H. (Eds.), El Bosque seco tropical en Colombia. Instituto de Investigación de recursos Biológicos Alexander von Humboldt, Bogota D.C. Colombia, pp. 229–251.
Goudet, J., 1999. FSTAT, a Program for IBM PC Compatibles to Calculate Weir and Cockerham's (1984) Estimator of F-Statistics (Version 1.2). University of Lausanne, Switzerland.
Hengl, T., de Jesus, J.M., MacMillan, R.A., Batjes, N.H., Heuvelink, G.B.M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Ruiperez Gonzalez, M., 2014. SoilGrids1 km — global soil information based on automated mapping. PLoS One 9, e105992.
Hewitt, G.M., 2004. Genetic consequences of climatic oscillations in the Quaternary. Philos. Trans. R. Soc. B Biol. Sci. 359, 183–195.
Hijmans, R.J., 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. Ecology 93, 679–688.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.
Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801–1806.
Kageyama, P.Y., Sebbenn, A.M., Ribas, L.A., Gandara, F.B., Castellen, M., Perecim, M.B., Vencovsky, R., 2003. Diversidade genética em espécies arbóreas tropicais de diferentes estágios sucessionais por marcadores genéticos. Sci. For. 64, 93–103.
Khurana, E., Singh, J.S., 2001. Ecology of seed and seedling growth for conservation and restoration of tropical dry forest: a review. Environ. Conserv. 28, 39–52.
Kindt, R., Coe, R., 2005. Tree Diversity Analysis. A Manual and Software for Common Statistical Methods for Ecological Biodiversity Studies. World Agroforestry Centre (ICRAF), Nairobi, Kenya.
Larranaga, N., Albertazzi, F.J., Fontecha, G., Palmieri, M., Rainer, H., Van Zonneveld, M., Hormaza, J.I., 2017. A Mesoamerican origin of cherimoya (*Annona cherimola* Mill.). Implications for the conservation of plant genetic resources. Mol. Ecol. 26, 4116–4130.
Lobo, J.A., Quesada, M., Stoner, K.E., 2005. Effects of pollination by bats on the mating system of *Ceiba pentandra* (Bombacaceae) populations in two tropical life zones in Costa Rica. Am. J. Bot. 92, 370–376.
Lowe, A., Cavers, S., Boshier, D., Breed, M., Hollingsworth, P., 2015. The resilience of forest fragmentation genetics—no longer a paradox—we were just looking in the wrong place. Heredity 115, 97–99.
Marchelli, P., Thomas, E., Azpilicueta, M.M., van Zonneveld, M., Gallo, L., 2017. Integrating genetics and suitability modelling to bolster climate change adaptation planning in Patagonian Nothofagus forests. Tree Genet. Genomes 13, 119.
Mayle, F.E., 2004. Assessment of the Neotropical dry forest refugia hypothesis in the light of palaeoecological data and vegetation model simulations. J. Quat. Sci. 19, 713–720.
de Melo, A.F., de Carvalho, D., Vieira, F.A., de Oliveira, D.A., 2012. Spatial genetic structure in natural populations of *Caryocar brasiliense* Camb.(Caryocareceae) in the North of Minas Gerais, Brazil. Biochem. Syst. Ecol. 43, 205–209.
de Melo, W.A., Lima-Ribeiro, M.S., Terribile, L.C., Collevatti, R.G., 2016. Coalescent simulation and paleodistribution modeling for *Tabebuia rosealba* do not support South American dry forest refugia hypothesis. PLoS One 11, e0159314.
Miles, L., Newton, A.C., DeFries, R.S., Ravilious, C., May, I., Blyth, S., Kapos, K., Gordon, J.E., 2006. A global overview of the conservation status of tropical dry forests. J. Biogeogr. 33, 491–505.
Mogni, V.Y., Oakley, L.J., Prado, D.E., 2015. The distribution of woody legumes in neotropical dry forests: the Pleistocene arc theory 20 years on. Edinb. J. Bot. 72, 35–60.
Moreira, P.A., Fernandes, G.W., Collevatti, R.G., 2009. Fragmentation and spatial genetic structure in *Tabebuia ochracea* (Bignoniaceae) a seasonally dry Neotropical tree. For. Ecol. Manag. 258, 2690–2695.
Moreira, P.A., Brandão, M.M., de Araujo, N.H., de Oliveira, D.A., Fernandes, G.W., 2015. Genetic diversity and structure of the tree *Enterolobium contortisiliquum* (Fabaceae) associated with remnants of a seasonally dry tropical forest. Flora Morphol. Distrib. Funct. Ecol. Plants 210, 40–46.
Motamayor, J.C., Lachenaud, P., e Mota, J.W.D.S., Loor, R., Kuhn, D.N., Brown, J.S., Schnell, R.J., 2008. Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). PLoS One 3, e3311.
Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89, 583–590.
Novaes, R.M.L., Rodrigues, J.G., Lovato, M.B., 2009. An efficient protocol for tissue sampling and DNA isolation from the stem bark of Leguminosae trees. Genet. Mol. Res. 8, 86–96.
Paloma-Lozano, E., 1994. Investigaciones forestales: resultados de dos ensayos de investigación, con 12 especies cada uno en suelos degradados. Bucaramanga, Colombia. Ministerio del Medio Ambiente, INDERENA-Regional Santander 40 pp.. http://documentacion.ideam.gov.co/cgi-bin/koha/opac-detail.pl?biblionumber=9853.
Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics 28, 2537–2539.
Pennington, R.T., Prado, D.E., Pendry, C.A., 2000. Neotropical seasonally dry forests and Quaternary vegetation changes. J. Biogeogr. 27, 261–273.
Petit, R.J., Aguinagalde, I., de Beaulieu, J.-L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Müller-Starck, G., Demesure-Musch, B., Palmé, A., Martín, J.P., Rendell, S., Vendramin, G.G., 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. Science 80, 1563–1565.
Pizano, C., García, H. (Eds.), 2014. El bosque seco tropical en Colombia. Instituto Alexander von Humbolt (IAvH), Bogotá D.C., pp. 353.
PNNC (Parques Nacionales Naturales de Colombia), 2018. Portafolio de nuevas áreas protegidas. Acceso 02 de enero de 2018. http://www.parquesnacionales.gov.co/portal/es/sistema-nacional-de-areas-protegidas-sinap/portafolio-de-nuevas-areas-protegidas-del-sistemas-de-parques-nacionales/.
Prado, D., Gibbs, P., 1993. Patterns of species distributions in the dry seasonal forests of South America. Ann. Missouri Bot. Gard. 80, 902–927.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
Prober, S.M., Byrne, M., McLean, E.H., Steane, D.A., Potts, B.M., Vaillancourt, R.E., Stock, W.D., 2015. Climate-adjusted provenancing: a strategy for climate-resilient ecological restoration. Front. Ecol. Evol. 3, 65.
Quesada, M., Stoner, K.E., Rosas-Guerrero, V., Palacios-Guevara, C., Lobo, J.A., 2003. Effects of habitat disruption on the activity of nectarivorous bats (Chiroptera: Phyllostomidae) in a dry tropical forest: implications for the reproductive success of the neotropical tree *Ceiba grandiflora*. Oecologia 135, 400–406.
Quesada, M., Stoner, K.E., Lobo, J.A., Herrerias-Diego, Y., Palacios-Guevara, C., Munguía-Rosas, M.A., O-Salazar, K.A., Rosas-Guerrero, V., 2004. Effects of forest fragmentation on pollinator activity and consequences for plant reproductive success and mating patterns in bat-pollinated Bombacaceous trees. Biotropica 36, 131–138.
Ramirez Villegas, J., Jarvis, A., 2010. Downscaling Global Circulation Model Outputs: The Delta Method. Decision and Policy Analysis Working Paper No. 1. Decision and Policy Analysis Working Papers. International Center for Tropical Agriculture (CIAT), Cali, Colombia.
Rosenberg, N.A., 2004. DISTRUCT: a program for the graphical display of population structure. Mol. Ecol. Notes 4, 137–138.
Rousset, F., 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genet. Mol. Biol. 145, 1219–1228.
Sánchez-Azofeifa, G.A., Kalacska, M., Quesada, M., Calvo-Alvarado, J.C., Nassar, J.M., Rodríguez, J.P., 2005. Need for integrated research for a sustainable future in tropical dry forests. Conserv. Biol. 19, 285–286.
Sánchez-Azofeifa, A., Powers, J.S., Fernandes, G.W., Quesada, M. (Eds.), 2013. Tropical Dry Forests in the Americas: Ecology, Conservation, and Management. CRC Press.
Sarmiento, G., 1975. The dry plant formations of South America and their floristic connections. J. Biogeogr. 233–251.
Schuelke, M., 2000. An economic method for the fluorescent labeling of PCR fragments. Nat. Biotechnol. 18, 233–234.
Sgro, C.M., Lowe, A.J., Hoffmann, A.A., 2011. Building evolutionary resilience for conserving biodiversity under climate change. Evol. Appl. 4, 326–337.
Sorngmenenye-Abengmeneng, C., Ofori, D., Kumapley, P., Akromah, R., Jamnadass, R., Quain, M., 2016. Genetic relationship among 36 genotypes of *Ceiba pentandra* (L.) as revealed by RAPD and ISSR markers. Am. J. Agric. Forest. 4, 86–96.
Thomas, E., Jalonen, R., Loo, J., Boshier, D., Gallo, L., Cavers, S., Bordács, S., Smith, P., Bozzano, M., 2014. Genetic considerations in ecosystem restoration using native tree species. For. Ecol. Manag. 333, 66–75.
Thomas, E., Alcázar Caicedo, C., McMichael, C.H., Corvera, R., Loo, J., 2015. Uncovering spatial patterns in the natural and human history of Brazil nut (*Bertholletia excelsa*) across the Amazon Basin. J. Biogeogr. 42, 1367–1382.
Thomas, E., Alcazar, C., Moscoso, H., L., G., Vasquez, A., Osorio, L.F., Salgado-Negret, B., González, M.A., Parra, M., Bozzano, M., Loo, J., Jalonen, R., Ramirez, W., 2017a. The importance of species selection and seed sourcing in forest restoration for enhancing adaptive potential to climate change: Colombian tropical dry forest as a model. CBD Tech. Ser. 89, 122–132.
Thomas, E., Gil Tobón, C., Gutiérrez, J.P., Alcázar Caicedo, C., Moscoso Higuita, L.G., Becerra, L.A., Loo, J., González, M.A., 2017b. Genetic diversity of *Enterolobium cyclocarpum* in Colombian seasonally dry tropical forest: implications for conservation and restoration. Biodivers. Conserv. 26, 825–842.
Thomas, E., Tovar, E., Villafañe, C., Bocanegra, J.L., Moreno, R., 2017c. Distribution, genetic diversity and potential spatiotemporal scale of alien gene flow in crop wild relatives of rice (*Oryza* spp.) in Colombia. Rice 10, 13.
Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P., 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes 4, 535–538.
Vergara, V.H., 2015. Patrones de la vegetación y tipos de uso de la tierra en el valle del Patía. Colomb. For. 18, 25–45.
Vieira-Blanco, A.J., Pereira, M.F., Siqueira-Guedes, A., Chaves, L.J., 2007. Diversidade genética em populações naturais de araticunzeiro (*Annona crassiflora* Mart.) por meio da análise de seqüências de cpDNA. Pesq. Agropec. Trop. 37, 169–175.
Vina, A., Cavelier, J., 1999. Deforestation rates (1938–1988) of tropical lowland forests on the andean foothills of Colombia. Biotropica 31, 31–36.
Vitorino, L.C., Lima-Ribeiro, M.S., Terribile, L.C., Collevatti, R.G., 2016. Demographical history and palaeodistribution modelling show range shift towards Amazon Basin for a Neotropical tree species in the LGM. BMC Evol. Biol. 16, 213.
Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370.
Zotz, G., Winter, K., 1994. Photosynthesis of a tropical canopy tree, *Ceiba pentandra*, in a lowland forest in Panama. Tree Physiol. 14, 1291–1301.